Journal of Medical Virology 59:437–450 (1999)

Artificial NS4 Mosaic Antigen of Hepatitis C Virus

J.C. Chang,[1] B. Ruedinger,[1] M. Cong,[1] S. Lambert ,[1] E. Lopareva,[1] M. Purdy,[1] B.P. Holloway,[2] D.L. Jue,[2] B. Ofenloch, [3] H.A. Fields,[1] and Y.E. Khudyakov[1]*

[1]*Hepatitis Branch, Division of Viral and Rickettsial Diseases, National Center for Infectious Diseases, Centers for Disease Control and Prevention, Public Health Service, U. S. Department of Health and Human Services, Atlanta, Georgia*

[2]*Biotechnology Core Facility Branch, Scientific Resources, National Center for Infectious Diseases, Centers for Disease Control and Prevention, Public Health Service, U. S. Department of Health and Human Services, Atlanta, Georgia*

[3]*Boehringer GmbH, D-82372 Penzberg, Germany*

An artificial antigen composed of 17 small antigenic regions derived from the NS4-protein of hepatitis C virus (HCV) genotypes 1 through 5 was designed and constructed. Eleven antigenic regions were derived from the 5-1-1 region, and 6 others were derived from the C-terminus of the NS4-protein of different genotypes. The gene encoding for this artificial antigen was assembled from synthetic oligonucleotides by a new approach designated as restriction enzyme-assisted ligation (REAL). The full-length synthetic gene was expressed in *Escherichia coli* as a fusion protein with glutathione S-transferase. By the use of site-specific antibodies raised against synthetic peptides, it was shown that all regions for which sequence-specific antibodies were obtained were accessible to antibody binding. The diagnostic relevance of the NS4 artificial antigen was demonstrated by testing this antigen with 4 HCV seroconversion panels and a panel of previously tested and stored serum specimens. The artificial antigen was found to specifically detect anti-NS4 antibodies in a number of specimens that were previously found to be anti-NS4 negative. Furthermore, this antigen detected anti-NS4 activity earlier in 2 of 4 seroconversion panels than did the antigen used in a commercially available supplemental assay. Equally important is the observation that the artificial NS4 antigen demonstrated equivalent anti-NS4 immunoreactivity with serum specimens obtained from patients infected with different HCV genotypes, whereas the NS4 recombinant protein derived from genotype 1, used in the commercial supplemental test, was less immunoreactive with serum specimens containing HCV genotypes 2, 3, and 4. Collectively, these data support the significant diagnostic potential of the NS4 mosaic antigen. The strategy employed in this study may be applied to the design and construction of other artificial antigens with improved diagnostically pertinent properties. ***J. Med. Virol. 59:437–450 1999.***

KEY WORDS: HCV; recombinant proteins; antigens; peptides; sequence-specific antibodies

INTRODUCTION

Hepatitis C virus (HCV), the main causative agent of non-A non-B hepatitis [Choo et al., 1989; Kuo et al., 1989], is classified as a new member of the *Flaviviridae* family. The HCV genome is a single, positive-strand RNA of approximately 9500 nucleotides containing one long open reading frame that encodes for a polyprotein of approx. 3000 amino acids (aa) [Choo, 1989; Kato, 1990; Choo, 1991; Takamizawa, 1991]. The N-terminal part of the polyprotein is processed into structural proteins, namely, nucleocapsid or core protein and two envelope proteins, E1 and E2. The C-terminal part is processed into nonstructural proteins, namely, NS2, NS3, NS4a, NS4b, and NS5a and NS5b proteins [Miller, 1990; Takeuchi, 1990; Choo, 1991; Hijikata, 1991; Houghton, 1991; Takamizawa, 1991; Grakoui, 1993].

The HCV genome is very heterogeneous [Dusheiko et al., 1994; Bukh et al., 1995; Simmonds, 1995b]. However, sequence heterogeneity is not uniformly distributed across the entire genome. The 5′-terminal part of the HCV genome, including the 5′-noncoding and core regions, is the most conserved, whereas E1, E2, NS4, and NS5a regions are very variable [Okamoto 1992;

Dr. Lambert's current address is Covance Central Laboratory Services Inc., Indianapolis, IN 46214–2985.

*Correspondence to: Dr. Y. E. Khudyakov, Hepatitis Branch, MS A-33, Division of Viral and Rickettsial Diseases, Centers for Disease Control and Prevention, 1600 Clifton Rd., Atlanta, GA 30333. E-mail: yek0@cdc.gov

Accepted 25 February 1999

Bukh, 1993; Simmonds, 1993a]. In particular, the N-terminal small region of the E2 gene, designated as hypervariable region 1 (HVR1), was found to be especially variable [Weiner, 1991; Hijikata, 1991]. The NS2, NS3, and NS5B genes are less variable than the E1, E2, NS4, and NS5a genes [Okamoto, 1992].

HCV heterogeneity has been extensively studied [Okamoto , 1992; Simmonds, 1993a; Buhk, 1995] resulting in the classification of HCV into six major genotypes [Simmonds, 1993a]. The effect of sequence heterogeneity on the antigenic properties of different HCV proteins was investigated in several studies [Chan et al., 1991; Machida et al., 1992; McOmish et al., 1993; 1994; Nagayama et al., 1993; Simmonds et al., 1993b; Alonso et al., 1994; Bhattacherjee et al., 1995; Zein, 1995; 1997; Dhaliwal et al., 1996]. A number of diagnostically relevant antigenic epitopes have been found within the core region [Muraiso et al., 1990; Chiba et al., 1991; Hosein et al., 1991; Nasoff et al., 1991; Sallberg et al., 1992], E1/E2 [Lok et al., 1993; Selby et al., 1993; Ray et al., 1994;], NS3 [Chien et al., 1992;], NS4 [Kuo et al., 1989; Chien et al., 1992;] and NS5 proteins [Mori, 1992; Riezu-Boj, 1992]. The HCV core, NS3 and NS4 proteins are the most diagnostically relevant antigens now used in various commercial immunoassays [Muraiso et al., 1990; Hosein et al., 1991; Van Der Poel et al., 1991; Chien et al., 1992; Inoue et al., 1992; Mori et al., 1992; Riezu-Boj et al., 1992; Berasain et al., 1993; Bukh, 1995]. The NS4 recombinant protein c100-3 of genotype 1 was shown less immunoreactive or nonimmunoreactive with antibodies in the serum specimens from the patients infected with HCV genotypes 2 and 3 compared to the immunoreactivity with the serum specimens from patients infected with genotype 1 [McOmish et al., 1993]. Genotype-specific antigenic reactivity was also observed for antigenic epitopes derived from core and NS4 proteins [Machida, 1992; Simmonds et al., 1993b, 1995a; Bhattacherjee et al., 1995]. The genotype-specific antigenic reactivity associated with diagnostically relevant HCV antigens has potentially serious diagnostic implications. Since all commercially available serological assays are based on antigens derived from genotype 1, the efficacy of detection of antibodies to different HCV genotypes may vary significantly. Accordingly, several publications have shown that commercial immunoassays have reduced immunoreactivity of the serum specimens from the patients infected with HCV genotypes 2, 3, 4, 5, and 6, compared to that from the patients infected with genotypes 1a or 1b [Simmonds et al., 1993b, 1995a; Bhattacherjee et al., 1995; Yuki, 1995; Zein, 1995, 1997; Dhaliwal, 1996].

The NS4 protein, which is the main focus of the present study, contains two strongly and broadly immunoreactive regions. One, designated as region 5-1-1, comprises the cleavage site separating NS4a and NS4b proteins [Grakoui et al., 1993]. The other region, designated hereafter as region 59, was found at the C-terminus of the NS4b protein [Khudyakov et al., 1995]. The primary structure of region 5-1-1 is very heterogeneous [Dusheiko et al., 1994; Bukh et al., 1995; Simmonds, 1995b]. This heterogeneity affects the antigenic properties of this region [McOmish, 1993; Simmonds, 1993b, 1995a; Zein, 1995, 1997; Dhaliwal, 1996] to such an extent that peptides derived from different HCV genotypes have been used to develop an assay that discriminates between different HCV serotypes [Simmonds et al., 1993b, Bhattachejee et al., 1995; Zhang et al., 1995]. Region 59 is less heterogeneous [Stuyver et al., 1994; Adams et al., 1997; Chamberlain et al., 1997]. However, a few amino acid substitutions within this sequence were shown to be essential for the antigenic epitope(s) 59 [Khudyakov, 1995] and suggest that variation in the primary structure may have an effect on the antigenic properties of this region as well.

The present paper describes a new, artificial NS4 antigen composed of antigenic epitopes derived from regions 5-1-1 and 59 from several HCV genotypes. The strategy of making artificial mosaic antigens has been previously described for the hepatitis E virus [Khudyakov et al., 1994]. The results obtained in the current study demonstrate that the artificial NS4 mosaic antigen detects anti-NS4 activity with greater sensitivity than the NS4 antigen derived from a single genotype 1 currently used in various commercial immunoassays.

MATERIALS AND METHODS

HCV Sequences

GenBank accession numbers for the amino acid sequences of region 5-1-1 (1691–1733 aa) used in this study are as follows: D10934, M86766, U14259, U14269, D14114, D14853, D14600, L29579, and D10988. The accession numbers for the amino acid sequences of region 59 (1921–1940) used in this study are as follows: D49374, D28917, U01214, D00944, D10988, and U16362.

Synthetic Oligonucleotides

Oligonucleotides were synthesized with an automatic synthesizer (model 380A; Applied Biosystems, Foster City, CA) and purified by either high-performance liquid chromatography or electrophoresis in a 10% polyacrylamide gel containing 7 M urea in Tris-borate-EDTA (TBE) buffer (0.045 M Tris-borate, 0.001 M EDTA, pH 8.3). The oligonucleotides were recovered from the gel by electroelution using a model 230A HPEC system (Applied Biosystems) according to the manufacturer's protocols.

Construction of Vector pCV3

First, 1 μg of pGEX-4T-2 (Pharmacia Biotech Inc., Piscataway, NJ) was incubated with 10 units of BamHI and 10 units of NotI in buffer E (Promega, Madison, WI) at 37°C for 6 hours followed by the addition of 0.3 pmol of dGTP, 0.3 pmol of dATP, and 5 units of Klenow fragment (Promega) and incubated for 30 minutes at room temperature. After phenol extraction, DNA was precipitated with ethanol. Finally, the treated vector was ligated with 50 pmol of two complementary oligonucleotides (5′- TCG CAG CGA ATT CTC GAG GAT

CCA TCC and 5′- CCG GAT GGA TCC TCG AGA ATT CGC TGC) by using the rapid DNA ligation kit (Boehringer Mannheim Biochemicals, Indianapolis, IN) according to the manufacturer's protocols.

Synthetic Gene Assembly

The synthetic gene was assembled from nine segments. Each segment was assembled from two oligonucleotides that had complementary sequences at the 3′ ends; 50 pmol of each oligonucleotide was used to start the gene assembly. After each pair of oligonucleotides was annealed by its 3′ complementary sequences, each mixture was treated with 5 units of Klenow fragment (Promega) in the presence of 1 mM dNTPs (Boehringer Mannheim) at room temperature for 30 minutes to convert single-stranded regions of each pair into double-stranded DNA fragments. Then, 1 μg of each DNA fragment was digested with 10 units of EcoRI and 10 units of BamHI (Boehringer Mannheim) for 6 hours or overnight at 37°C. These fragments were purified from the incubation mixture by using the Wizard PCR Prep DNA purification system (Promega), according to the manufacturer's protocols, and ligated with the vector pCV3, which was preliminarily cleaved with same enzymes, by using the rapid DNA ligation kit (Boehringer Mannheim). The *Escherichia coli* competent cells JM109 (Promega) were transformed with the ligation mixtures according to the provided protocol (Boehringer Mannheim). Recombinant plasmids were recovered from transformants by using the Wizard Miniprep DNA Purification System (Promega). The presence of an insert was confirmed by PCR using two primers of the following sequences: 5′- GTA CTT GAA ATC CAG CA and 5′- GTT TTC ACC GTC ATC AC, derived from regions flanking the multiple cloning site of pGEX-4T-2. The primary structure of inserts was confirmed by DNA sequencing. PCR fragments of the correct size and primary structure were treated with 24 units of BbvI (Promega) overnight at 50°C or 18 units of Fok I (Promega) overnight at 37°C. The digested fragments were purified by using the Wizard PCR Prep DNA purification system (Promega). Every two consecutive PCR fragments, one digested with BbvI and the other digested with FokI, were ligated together by using the rapid DNA ligation kit (Boehringer Mannheim), amplified by using the same primers derived from pGEX-4T-2 as mentioned above, and cloned with the vector pCV3. This protocol of ligation and cloning of two consecutive DNA fragments was repeated again with each extended fragments until a full-length gene was assembled.

DNA Sequencing

Sequencing was performed by using an automated sequencer (373 DNA sequencer; Applied Biosystems) according to the manufacturer's protocol.

Gene Expression and Protein Purification

E. coli JM109 cells transformed with recombinant plasmids were grown overnight in Luria broth containing 50 μg/ml ampicillin at 37°C. The overnight culture was diluted 20 times with fresh Luria broth containing 50 μg/ml ampicillin and grown for 3 to 4 hours until an optical density (OD) value of 0.6–1.0 at 600 nm was reached. The gene expression was activated by the addition of isopropyl-β-D-thiogalactopyronoside (IPTG, Sigma Chemical Co., St. Louis, MO) at a final concentration of 1 mM. After 1 hour of growth at 30°C, cells were harvested and lysates were prepared [Sambrook et al., 1989]. The glutathione S-transferase (GST) mosaic proteins were purified by ligand affinity chromatography [Smith & Johnson, 1988] by using glutathione sepharose 4B column (Pharmacia Biotech, Piscataway, NJ).

Immunoblot Assay

Aliquots of each lysate or aliquots of the purified GST mosaic proteins were subjected to electrophoresis on precast 12% sodium dodecyl sulfate-polyacrylamide gels (Bio-Rad, Richmond, CA) followed by blotting onto a nitrocellulose membrane (Bio-Rad). Following protein transfer, the nitrocellulose membranes were incubated with blocking solution (0.1 M phosphate-buffered saline (PBS) containing 1% bovine serum albumin (BSA), 0.5% Tween 20, and 10% normal goat serum) overnight at 4°C, and then incubated with human anti-HCV positive serum specimens diluted 1:100 or 1:200 in blocking solution for 1 hour at room temperature. For immunodetection, membranes were washed three times with blocking solution, followed by the addition of affinity-purified goat anti-human immunoglobulin (IgG) conjugated to horseradish peroxidase (HRP), diluted 1:4000 or 1:6000 in blocking solution, and incubated for 1 hour at room temperature. After the membranes were washed with blocking solution three times, diaminobenzidine (Sigma) and hydrogen peroxide were added to develop the reaction.

Enzyme Immunoassay for Anti-HCV

For the detection of anti-HCV activity in serum specimens with recombinant proteins, 110 μl of protein was adsorbed onto microtiter wells (Immuno II; Dynatech Industies, Chantilly, VA) at a concentration of 100 ng/ml in 0.1 M PBS, pH 7.5, overnight at room temperature. The wells were washed with PBS/Tween buffer (0.1 M PBS containing 0.05 % Tween 20) 5 times before incubating with human serum specimens. Serum specimens were diluted 1:500 in blocking solution (as described above for the immunoblot assay) and incubated in the microtiter wells with the preadsorbed recombinant protein for 1 hour at 37°C. After the microtiter wells were washed, goat anti-human IgG conjugated to horseradish perxidase (HRP) diluted 1:4000 times in blocking solution, was added to the wells and incubated for 1 hour at 37°C. The microtiter wells were washed again for 5 times, and then subjected to color development by adding *o*-phenylenediamine according to the provided protocol (Abbott Laboratories, North Chicago, IL). The reaction was stopped with 2 *N* sulfuric acid and the OD was measured at 493 nm. The

cutoff (CO) was statistically established as the mean of negative controls plus 3.5 standard deviations of the mean. The results were expressed as a ratio (S/CO) between the OD value of the tested specimen (S) and the OD value equal to CO. A S/CO value greater than 1 was considered as an indicator of positive immunoreactivity.

For the analysis of antisera, 1 μg of each synthetic peptide was absorbed onto the surface of microtiter wells. Antisera against irrelevant peptides were used as negative controls. Serum specimens were diluted 1:50 or 1:10 and goat anti-mouse antibody conjugated with HRP was used to detect binding between peptides and antipeptide antibodies. The results were expressed as a ratio (P/N) between the OD value of tested specimen (P) and the mean of negative controls (N). The CO established as a P/N value equal or greater than 5 times OD value of the mean of negative controls.

Human Sera

Anti-HCV-positive serum specimens ($n = 166$) were obtained from Boston Biomedical Inc. (West Bridgewater, MA) and Boehringer Mannheim GmbH, (Penzberg, Germany). A genotyped panel of serum specimens was obtained from a collection reposited at Boehringer GmbH Anti-HCV-negative serum specimens ($n = 160$) were obtained from a collection of normal human blood donors reposited at the Centers for Disease Control and Prevention (CDC, Atlanta, GA). The anti-HCV status of all serum specimens was confirmed with the second-generation anti-HCV EIA (Abbott) and a semiautomated dot blot immunoassay (MATRIX, Abbott). Serum specimens were also tested for HCV viral RNA using the AMPLICOR HCV MONITOR test, according to a commercially available protocol (Roche Diagnostic System, Branchburg, NJ). Serum specimens were also tested for HCV viral RNA by using the AMPLICOR HCV MONITOR test, according to the commercial available protocol (Roche).

HCV Seroconversion Panels

Three HCV seroconversion panels (4811, 4812, and 4813) were obtained from Serologicals (Clarkston, GA). The HCV seroconversion panel (6214) was obtained from BioClinical Partners (Franklin, MA). All samples from these panels were obtained as undiluted serum and not heat-inactivated.

Synthetic Peptides

Peptides were synthesized by FMOC chemistry [Barany & Merrifield, 1980] on an ACT model MPS 250 multiple peptide synthesizer (Advanced Chemtech, Louisville, KY), according to the manufacturer's protocols. The peptides were analyzed by amino acid analysis, high performance liquid chromatography, and capillary electrophoresis.

Mouse Antipeptide Sera

Each synthetic peptide (2 mg) was conjugated with 2 mg of BSA by 1-ethyl-3-(3-dimethylaminopropyl) carbodiimide hydrochloride (EDC) coupling methods by using a commercially available kit (Pierce, Rockford, IL). The adjuvant, Titer Max (CytRx, Atlanta, GA), was mixed with conjugated peptides [Aichele et al., 1990] before inoculation. Mice were inoculated subcutaneously at two sites on their back with 50 μl of the solution containing 25 μg of conjugated peptide per site. With the same mode of immunization, the animals were boosted 2 weeks later and bled 4 weeks later.

RESULTS

Protein Design

The NS4 mosaic protein was designed to contain antigenic epitopes from region 5-1-1 [Choo et al., 1989] and 59 [Khudyakov et al., 1995] from different HCV genotypes. Region 5-1-1 contains two distinct antigenic domains at position 1691–1710 aa and 1712–1733 aa [Khudyakov, 1995]. Five sequences derived from the first domain at position 1691–1710 aa with GenBank accession numbers D10934, M86766, U14259, U14269, and D14114, and six sequences derived from the second antigenic domain at position 1712–1733 aa with GenBank accession numbers D14853, D14600, L29579, D10988, D10934, and U14259 from HCV genotypes 1, 2, 3, and 5 were selected to be included into the mosaic protein (Fig.1). Six sequences derived from region 59 at position 1921–1940 aa with accession numbers D49374, D28917, U01214, D00944, D10988, and U16362 from HCV genotypes 1, 2, and 3 also were selected. In total, 17 regions were used to construct the NS4 mosaic protein (Fig. 1). No sequences from HCV genotypes 4 and 6 were used to design this protein because at the time of sequence selection these sequences were not available in GenBank. For region 5-1-1, the selection of sequences was based on identifying representative sequences that are closely related to the consensus sequence for each genotype. For region 59, all known different sequences were selected. All 17 regions were arranged within the NS4 mosaic protein in such an order that 1) regions would be scattered across the artificial protein, and 2) the predicted secondary structure for each individual region within this protein [Ptitsyn and Finkelstein, 1983] would be identical to the predicted secondary structure of the same region within the native HCV polyprotein.

Design and Assembly of Synthetic Gene

The amino acid sequence of the NS4 mosaic protein was reverse translated into DNA by using optimal *E. coli* codons (Fig. 1). The resulting sequence contained 1059 nucleotides. This nucleotide sequence was divided into 9 segments, with 8 segments encoding for 2 antigenic regions each and 1 segment encoding for only 1 antigenic region at the C-terminus of the NS4 mosaic protein (Fig. 1). Each segment was assembled from two oligonucleotides. One oligonucleotide was derived from the sense strand and the other was derived from the antisense strand of the gene. The 3′ ends of these oligonucleotides were designed to be complementary (Fig. 1). Each segment was designed in such a way that when oligonucleotides were annealed they could be converted

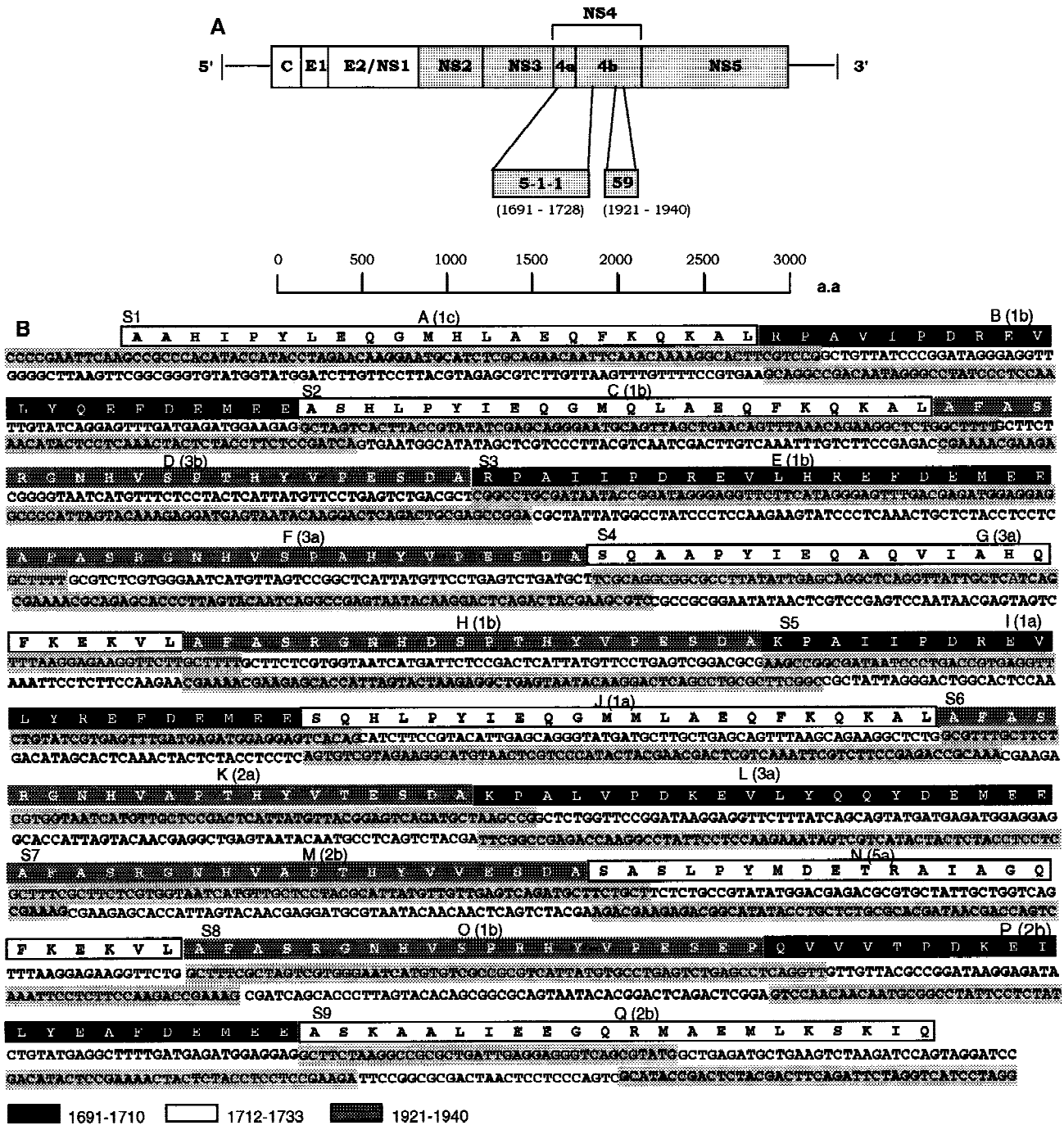

Fig. 1. **(A)** Schematic map of the HCV genome and amino acid position of the regions 5-1-1 and 59. **(B)** The amino acid sequence of the NS4 mosaic protein and the nucleotide sequence of the synthetic gene. Amino acid sequences derived from the region 5-1-1 antigenic domain at position 1691–1710 aa are shown in black boxes, and the antigenic domain at position 1712–1733 aa in white boxes. Amino acid sequences of region 59 are darkly shaded. Nucleotide sequences of synthetic oligodeoxyribonucleotides used in gene assembly are lightly shaded. S1 to S9 label the N-terminus of each segment. The HCV genotype and subtype of each antigenic region, A to Q, are shown in parentheses.

into double-stranded DNA by Klenow polymerase. The double-stranded DNA contained terminally located EcoRI and BamHI sites, which were used to clone these segments with pCV3 (Fig. 2) (see the Materials and Methods section). Additionally, each segment contained a TAG codon that was designed to stop translation at the end of each segment (Fig. 2).

After segments were inserted into pCV3, each segment was flanked with two sets of restriction endonuclease recognition sites (Fig. 2). On one side, it was flanked with BbvI and EcoRI sites. On the other side, it was flanked with BamHI and Fok I sites. Additionally, each segment, except for the 3′-terminal segment 9, contained 2 extra codons preceding TAG, which were

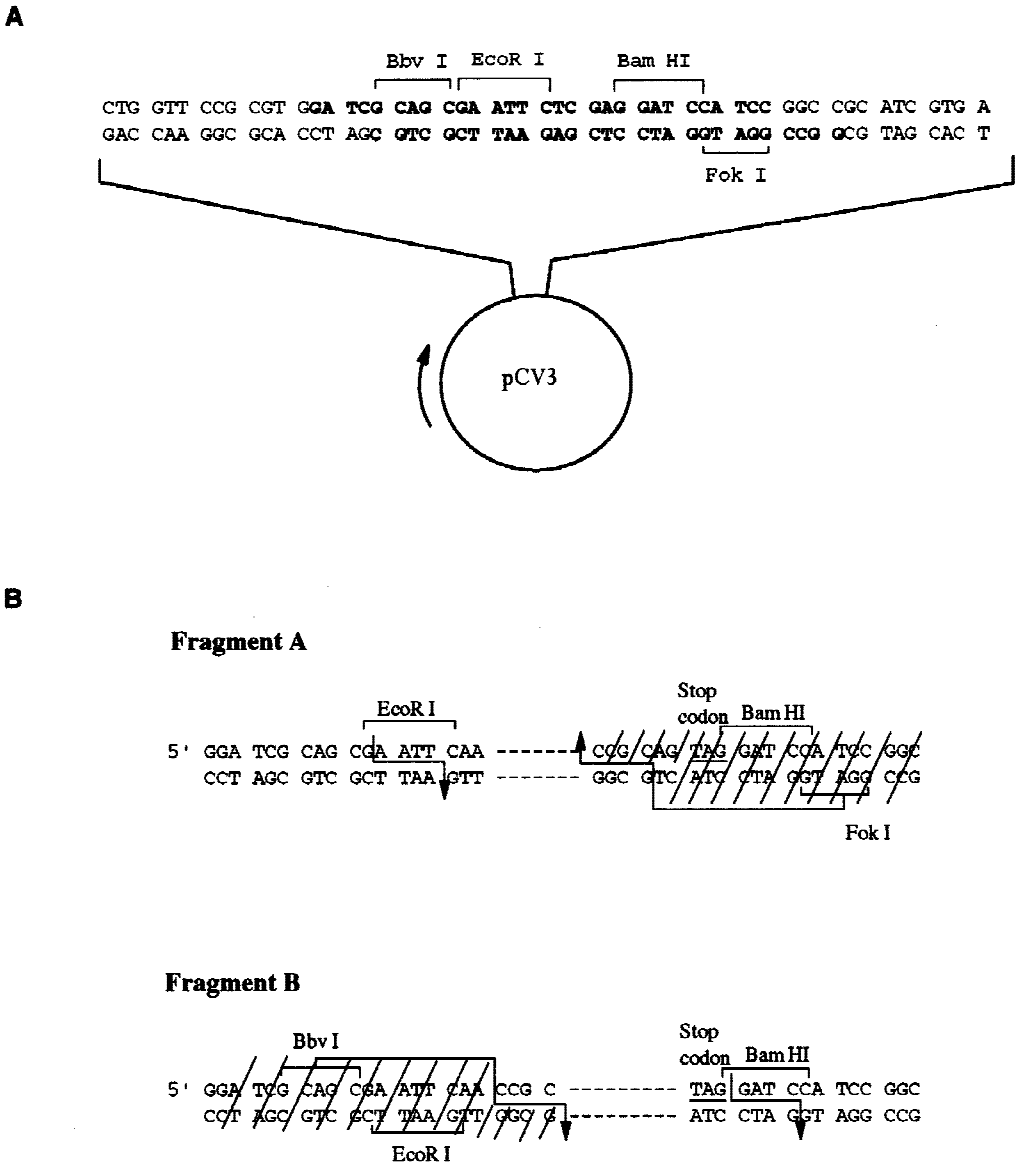

Fig. 2. **(A)** Multiple cloning site of the vector pCV3. **(B)** Design of fragments used for assembly of the synthetic gene by restriction enzyme assisted ligation (REAL). Arrows indicate the cleavage sites of restriction enzymes. Stop codon, TAG, is underlined. The sequences of each segment between restriction enzyme recognition sites are shown by dashed line. Sequences removed from the fragment A by FokI-cleavage and from the fragment B by BbvI-cleavage are shaded.

derived from the 3′- adjacent segment. For example, fragment A contains two codons, CCG and CAG, close to the BamHI site. Fragmant B contains the same two codons close to EcoRI site (Fig. 2B). When fragment A is cleaved with Fok I, this cleavage removes the TAG, stop codon and the BamHI site from this fragment, leaving a 5′ single-stranded protrusion with the sequence 5′-GGCG. When fragment B is cleaved with BbvI, this cleavage removes the EcoRI site from the fragment, leaving a 5′ single-stranded protrusion with the sequence 5′-CCGC, which is complementary to the 5′-protrusion of fragment A (Fig. 2B). Therefore, these two fragments can be ligated together into a dimer AB through these complementary single-stranded regions. This process of assembling double-stranded DNA from synthetic oligonucleotides through subsequent and repetitive cleavage with restriction enzymes and ligation was designated as restriction endonuclease assisted ligation (REAL).

After all nine segments were cloned, each fragment was amplified from the recombinant plasmid by using PCR primers derived from the vector sequences flanking the inserts (see the Materials and Methods section). PCR fragments containing segments S1 and S2, S3 and S4, S5 and S6, S7 and S8 were treated with FokI or BbvI and ligated together into dimers S1 + S2 (D1), S3 + S4 (D2), S5 + S6 (D3) and S7 + S8 (D4), as was described above for fragments A and B. Each dimer was amplified with the same pCV3-derived primers and, after being cleaved with EcoRI and BamHI, was

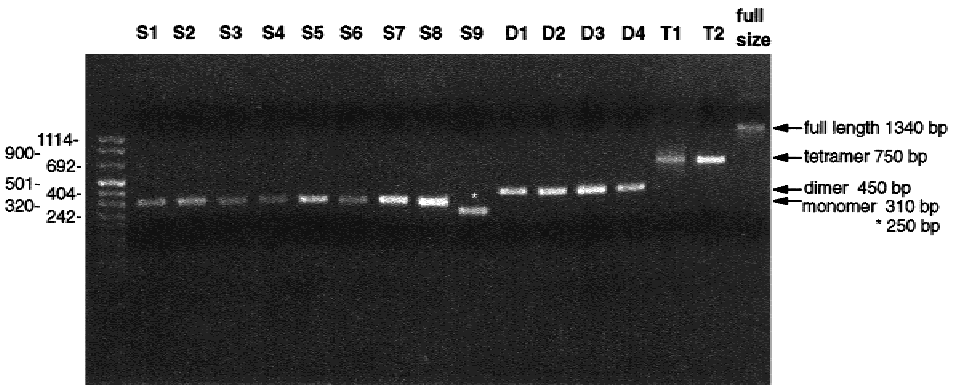

Fig. 3. Agarose gel electrophoresis analysis of PCR products obtained by amplification of monomers S1–S9, dimers D1–D4, tetramers T1 and T2, and the full-length gene.

inserted into pCV3 vector. Consecutive dimers were subsequently assembled into tetramers (T1 and T2), following the same strategy as for dimer assembly. Finally, two tetramers, T1 and T2, and segment S9 were assembled together into a full-length gene (Fig. 3). The primary structure of each individual segment, as well as the full-length gene, was confirmed by direct DNA sequencing.

Gene Expression

All nine segments, four dimers, two tetramers, and the full-length gene were expressed at high level as fusion proteins with glutathione S-transferase (Fig. 4). The level of expression diminished with the increasing length of the gene from monomers to dimers, or from tetramers to the full-length NS4 mosaic gene (Fig. 4). A similar trend was observed for solubility of the gene products. The larger proteins were increasingly less soluble. The full-length protein was almost entirely insoluble when expressed at 37°C (data not shown). To obtain this protein in soluble fraction, *E. coli* cells transformed with the plasmid containing the full-length mosaic gene were grown at a lower temperature of 30°C for no more than 1 hour after IPTG induction.

All proteins were purified by using ligand affinity chromatography (see the Materials and Methods section). The antigenic reactivity of these proteins was tested by immunoblot analysis with one serum specimen with high anti-NS4 activity. As can be seen in Figure 5, almost all proteins, except for proteins S1, S6, and S9, were strongly immunoreactive. Proteins S1 and S6 were found immunoreactive with the same serum specimen at a lower dilution (data not shown). Protein S9 was found weakly immunoreactive with several anti-NS4 positive serum specimens when tested by EIA (data not shown). Thus, all proteins containing different parts of the NS4 mosaic protein as well as the full-length protein were found immunoreactive. The difference in immunoreactivity among proteins S1, S6, and S9 and the other proteins can not be explained by any difference in the protein expression efficiency because all proteins from S1 to S9 were expressed at approximately the same level (Fig. 4). However, variations in specificity and quantity of anti-NS4 antibodies in different serum specimens may play an essential role in the efficiency of antibody binding by S proteins which contain divergent sequences from different NS4 antigenic regions derived from various HCV genotypes (Fig. 1B).

Immunoreactivity With Sequence-Specific Antibodies

After the mosaic protein was constructed, one of the most important issues is the accessibility of all antigenic epitopes to immunoreact with antibodies. To assess the accessibility of antigenic epitopes, sequence-specific antibodies were obtained that could recognize at least some individual antigenic regions within the NS4 mosaic protein. For this, 36 short peptides containing unique amino acid sequences specific for each of 17 antigenic regions were designed and synthesized (Table I). These peptides were conjugated with BSA and used to immunize mice (see the Materials and Methods section).

The specificity of anti-peptide antibodies was examined with 20-mer peptides containing sequences of all antigenic regions included in the NS4 mosaic protein [Chang et al., 1998]. These peptides were organized into 36 subsets. Each subset was composed of three peptides, one containing the sequence included in peptides designed for immunization at the N-terminus, another in the center, and the last peptide at the C-terminus. Seven mouse antisera against peptides 8, 10, 24, 26, 29, 31, 35 (Table I) showed strict sequence-specific immunoreactivity with peptides containing the sequences derived from antigenic regions G, I, L, P, F, and O within the NS4 mosaic protein (Fig. 1). Antiserum obtained against peptide 23 demonstrated strong immunoreactivity with a 20-mer peptide derived from region E (P/N = 102) and about 10- to 20-fold lower immunoreactivity with peptides derived from regions A, C, I, and J (Table I). Despite this lowered immunoreactivity to these other regions, this antiserum was also considered as sequence specific to region E. Antiserum to peptide 1 cross-immunoreacted with 20-mer peptides derived from regions A and C. Antiserum to peptide 5 cross-immunoreacted with 20-mer peptides

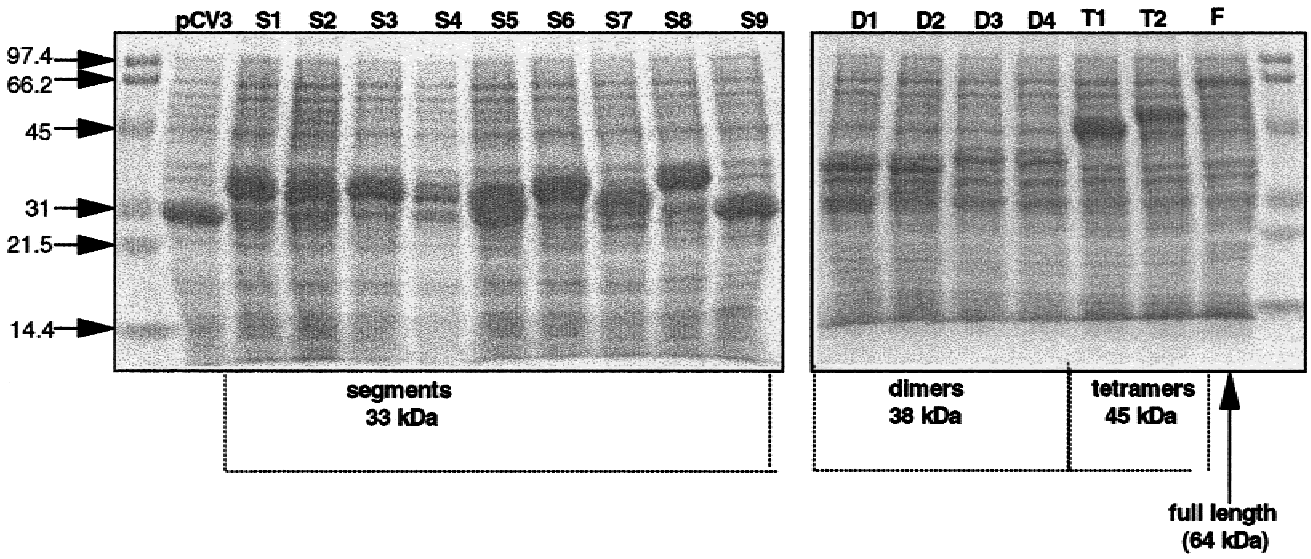

Fig. 4. Sodium dodecyl sulfate–polyacrylamide gel electrophoresis analysis of crude lysates prepared after IPTG einduction of *E. coli* transformed with plasmids containing monomers S1–S9, dimers D1–D4, tetramers T1 and T2, and full-length synthetic gene (F). The gel was Coomassie Blue stained.

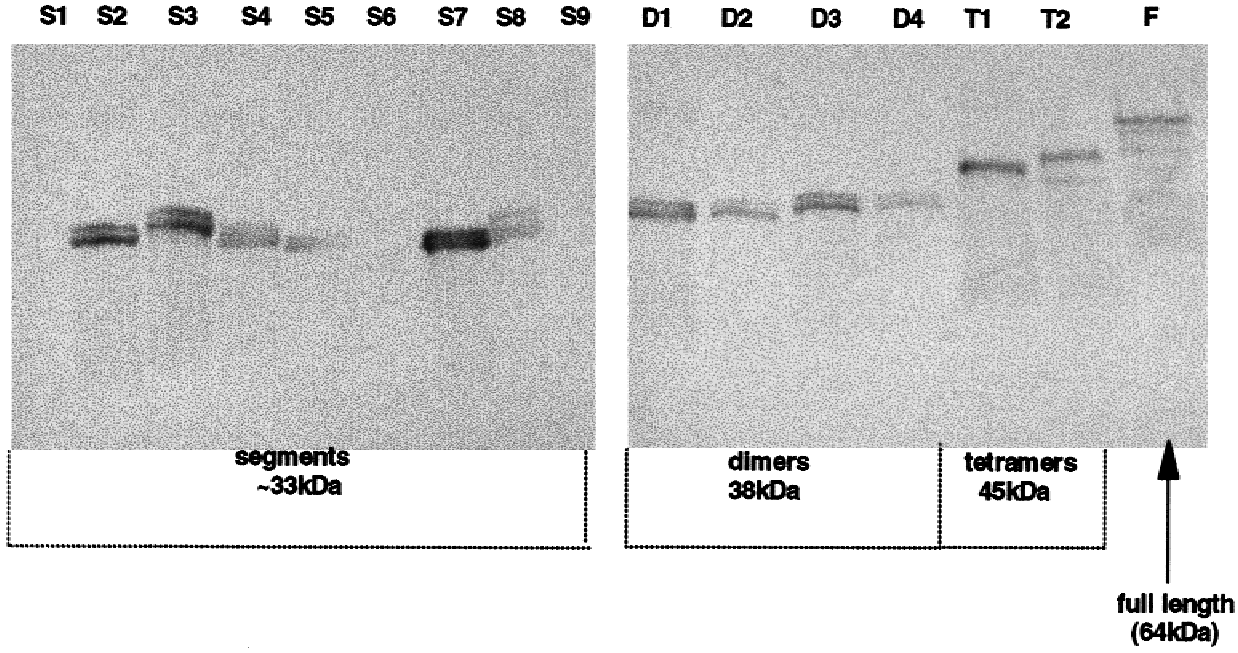

Fig. 5. Immunoblot analysis of purified proteins.

derived from 3 regions (A, C, and G), although showing preferential immunoreactivity with peptides from regions A and C compared to region G (Table I). Therefore, antisera to peptides 1 and 5 were considered sequence specific to 2 segments, A and C. The majority of antisera were not immunoreactive with 20-mer synthetic peptides (Table I), although all demonstrated strong immunoreactivity with the BSA-conjugated peptides used as immunogens. These data suggest that the lack of immunoreactivity with nonconjugated peptides is not due to the lack of an immunoresponse.

Antisera against peptide 8, 10, 23, and 29 were found to be immunoreactive with the NS4 protein when diluted at 1:50. Antisera against peptides 24 and 31 diluted at 1:10 were also shown to be immunoreactive with the NS4 mosaic protein (Table I). Thus, 6 out of 8 sequence-specific antisera demonstrated immunoreactivity with this artificial antigen. This finding strongly suggests that at least 5 regions (E, F, G, I, and P) are exposed to antibody binding. Antisera against peptides 26 and 35 demonstrated strong immunoreactivity with 20-mer synthetic peptides, but did not immunoreact

TABLE I. Antigenic Reactivity of the NS4 Mosaic Protein With Site-Specific Antibodies

Segment number	Region	Peptide number	Sequence of Peptides	Specificity of antisera 1:50 (P/N)[a]	Reactivity of Antisera to Mosaic Protein (P/N) 1:50[b]	1:10[b]
1	A	1	AAHIPYLEQG	A (76), C (20)	—	—
1	A	2	PYLEQG	—	—	—
1	A	3	HIPYLE	—	—	—
1	A	4	GMHLAE	—	—	—
2	C	5	ASHLPYI	A (55), C (47), G (6)	6.7	NT
2	C	6	GMQLAE	—	—	—
4	G	7	SQAAPYI	—	3.3	NT
4	G	8	EQAQVIAHQF	G (5)	7.1	NT
4	G	9	AQVIAH	—	—	—
4	G	10	AQVIAHQFKEKV	G (16)	4.1	NT
4	G	11	FKEKVL	—	—	—
5	J	12	SQHLPYI	—	—	—
5	J	13	GMMLAE	—	—	—
7	N	14	YMDETRAIAG	—	—	—
7	N	15	SASLPYMDET	Not reproducible	—	3.4
7	N	16	AIAGQFKEKV	—	10.3	NT
9	Q	17	ASKAALIEEGQ	—	—	—
9	Q	18	EGQRMAEMLK	—	—	—
9	Q	19	QRMAEMLKSKIQ	—	—	—
1	B	20	RPAVIPDR	—	—	3.0
1	B	21	LYQEFD	—	—	—
3	E	22	RPAIIPDR	—	3.94	NT
3	E	23	VLHREF	E (102, ACIJ (5-11)	10.0	NT
5	I	24	KPAIIPDR	I (5)	—	9.0
5	I	25	LYREFD	—	—	—
6	L	26	PALVPDKEV	L (91)	—	—
6	L	27	YQQYDE	—	—	—
8	P	28	QVVVTPDKEI	—	—	—
8	P	29	TPDKEILYEAFD	P (108)	7.9	NT
2	D	30	HVSPTH	—	—	—
3	F	31	SPAHYV	F (72)	—	5.5
4	H	32	HYVTES	—	—	—
6	K	33	HVAPTH	—	—	—
6	K	34	YVVESD	—	—	—
8	O	35	SPRHYV	O (240)	—	—
8	O	36	VPESEPGGGG	—	—	—

[a]Regions that antisera recognized.
[b]1:50 or 1:10—serum dilution.
—, No reaction; NT, Not tested; P36, GGGG are natural sequence of HCV; P/N, OD value of tested antisera over mean of the negative controls.
See figure 1b for segment number and Region.

with the NS4 mosaic protein (Table I). This finding may indicate that regions L and O are buried inside the protein globule and not exposed for antibody binding. However, we favor the alternative explanation that different synthetic peptides and proteins model antigenic epitopes in different ways. Thus, some antisera, like antisera to peptides 26 and 35, cannot recognize antigenic epitopes within the NS4 mosaic protein, whereas some other antisera, like antisera to peptides 7, 16, 20, and 22, can recognize the epitopes within the protein but not the same antigenic epitopes when modeled with peptides of different sizes (Table I).

Antiserum against peptide 5 also immunoreacted with the NS4 mosaic protein (Table I). Since these antisera showed preferential immunoreactivity with regions A and C, it is conceivable that at least one of these regions is accessible to antibodies. Antiserum to peptide 15 was weakly immunoreactive with its corresponding 20-mer peptide; however, this result was not reproducible. Nevertheless, these antisera were shown weakly immunoreactive with the NS4 mosaic protein (Table I). Thus, region N may be accessible to antibody binding as well. Collectively, the results of these experiments with antisera demonstrated that 5–8 antigenic regions scattered across the entire NS4 mosaic protein are exposed on the surface of the protein molecule in such a way as to be efficiently bound by antibodies.

Immunoreactivity With Human Serum Specimens

Specific immunoreactivity of the NS4 mosaic protein with anti-HCV antibodies was studied by using 160 anti-HCV-negative serum specimens from normal blood donors and 166 anti-HCV-positive serum specimens obtained from HCV-infected patients. The anti-HCV status of all serum specimens was initially confirmed by a second-generation screening test, a supple-

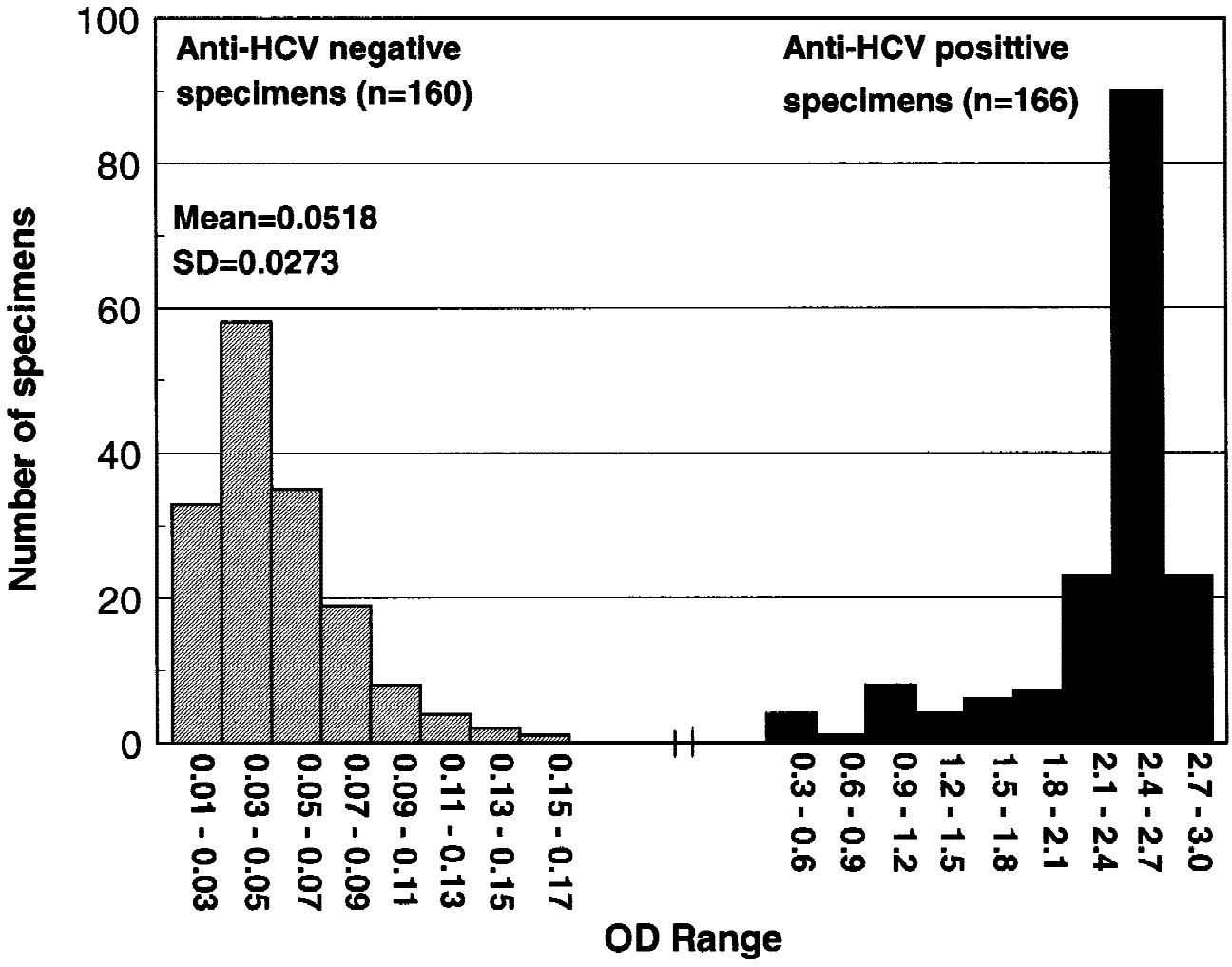

Fig. 6. Frequency distribution of OD values obtained by testing 160 anti-NS4-positive and 166 anti-HCV-negative serum specimens.

mental test and PCR (see the Materials and Methods section). The results showed that approximately 90% of anti-HCV-negative serum specimens had OD values less than 0.09, while approximately 80% of anti-HCV-positive sera had OD values greater than 2.1 (Fig. 6). The mean OD value for the anti-HCV-negative specimens was 0.052. A calculated CO value, the mean OD value of the anti-HCV-negative specimens plus 4.3 SD, clearly separated anti-HCV-negative from positive serum specimens (Fig. 6). Among these anti-HCV-positive serum specimens, 33 were not immunoreactive with the NS4-antigens in MATRIX (Abbott). However, anti-NS4 antibodies were detected with the NS4 mosaic protein in 25 of these 33 serum specimens, indicative of a very high clinical sensitivity for the artificial antigen in detecting anti-NS4 activity in serum specimens. All of these 25 specimens were either immunoreactive with at least 2 antigens in MATRIX or immunoreactive with only 1 antigen in MATRIX and HCV PCR positive, thus, confirming the HCV-positive status of these specimens.

Immunoreactivity With Serially Diluted Anti-HCV-Positive Serum Specimens

The sensitivity of the NS4 mosaic protein in detecting anti-NS4 antibodies was additionally examined by using two serially diluted anti-NS4 positive serum specimens. Each serial dilution was also tested by MATRIX (Abbott) for its anti-NS4 activity and demonstrated endpoint titers of 1:2000 and 1:4000, respectively. When the NS4 mosaic protein was used to test these serially diluted specimens, the endpoint titers were 1:32000 and 1:64000, respectively. Thus, the results of this experiment demonstrated a very high assay sensitivity for the detection of anti-NS4 activity by EIA when the NS4 mosaic protein was used as antigenic target.

Immunoreactivity With HCV Seroconversion Panels

Four HCV seroconversion panels were used in this study (see the Materials and Methods section). Anti-NS4 activity was first detected by MATRIX 63, 48, 53,

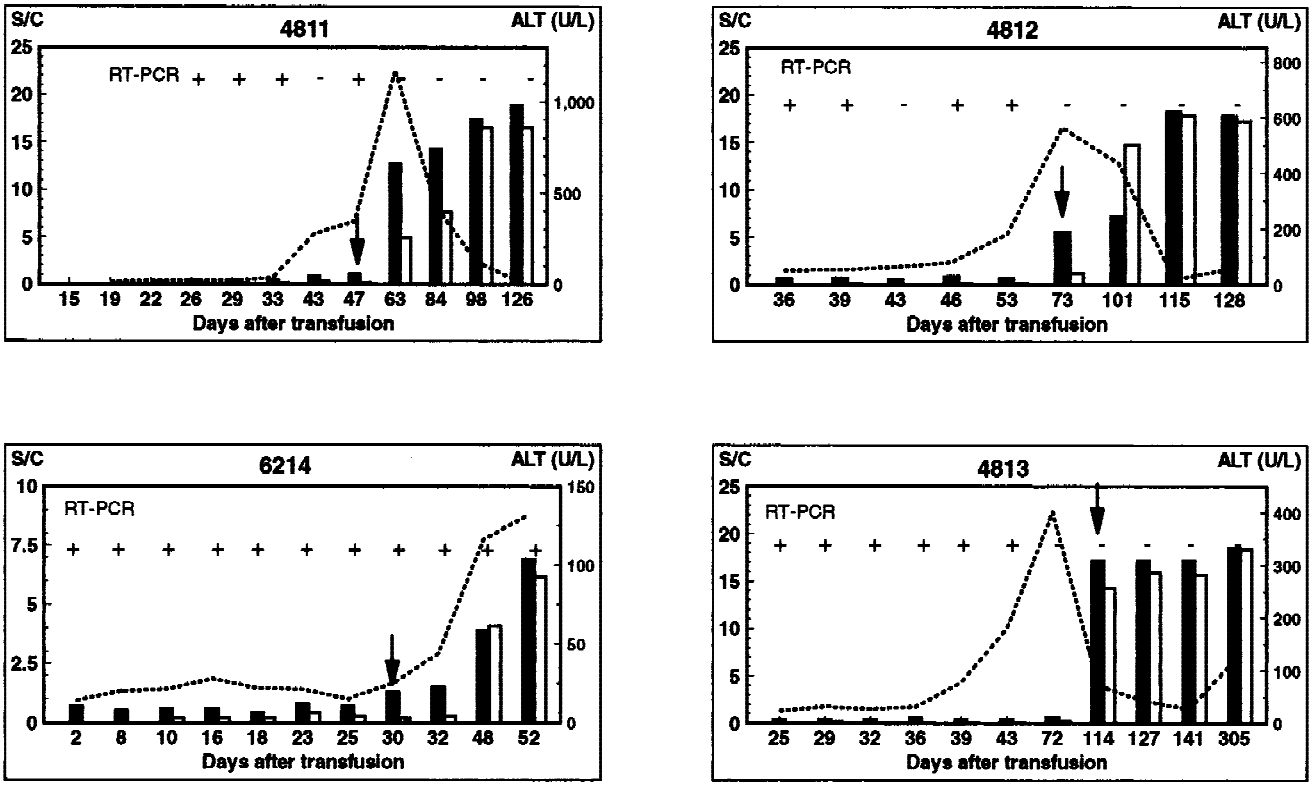

Fig. 7. Analysis of HCV seroconversion panels with the NS4 mosaic protein (black bars) and the NS4 protein (open bars) employed in MATRIX (Abbott Laboratories). ALT level is shown with dotted line. PCR results are shown by + (positive) and – (negative).

and 72 days after transfusion in panels 4811, 6214, 4812, and 4813, respectively (Fig. 7). When the same panels were tested with the NS4 mosaic protein, anti-NS4 activity was detected 47, 30, 53, and 72 days after transfusion, respectively (Fig. 7). Because some panels contain specimens separated by relatively large intervals, it is not always possible to make an exact estimate of any time difference between the detection of anti-NS4 antibody by the NS4 mosaic protein and by conventional NS4 recombinant antigens. Nevertheless, the results indicate that the NS4 mosaic protein detected anti-NS4 seroconversion in panels 4812 and 6214 several days earlier than the NS4 antigens included in MATRIX (Fig. 7).

Immunoreactivity With Antibodies to Different HCV Genotypes

The NS4 mosaic protein is composed of antigenic regions derived from several HCV genotypes. The rationale was to construct a protein that would immunoreact with anti-NS4 antibodies against different HCV genotypes with equal efficiency. Accordingly, this protein was tested with a panel of 22 serum specimens obtained from patients infected with HCV genotypes 1, 2, 3, 4, and 5 (Table II). These specimens had been tested previously by MATRIX to identify their anti-NS4 status. All genotype 1 specimens were anti-NS4-positive. Among genotype 2 and 3 specimens, only one specimen from each genotype group was anti-NS4-positive. One specimen was anti-NS4-negative among genotype 4 specimens, and 2 specimens were anti-NS4-negative among genotype 5 specimens (Table II). However, when these specimens were tested with the NS4

TABLE II. Antigenic Reactivity of NS4 Recombinant Protein and Mosaic Protein to Genotyped Serum Specimens (Signal/Cutoff value)

Genotype	Serum number	NS4 Antigen MATRIX (yeast)	NS4 Antigen MATRIX (*E. coli*)	NS4 Mosaic Protein
Type 1	102 104 01	14.0	15.6	6.5
	102 109 01	12.0	13.0	19.3
	102 110 01	16.1	17.6	17.3
	102 113 01	16.4	18.1	19.6
	102 114 01	10.5	12.1	7.6
Type 2	102 123 01	0.1	0.4	11.5
	101 43 01	0.2	0.5	0.2
	100 61 01	2.4	3.8	17.0
	100 43 01	0.1	0.5	17.8
Type 3	102 115 01	0.2	0.5	9.5
	101 27 01	0.2	0.3	0.2
	101 40 01	0.4	0.9	17.1
	101 38 01	0.1	0.6	15.3
	101 49 01	4.9	5.8	16.9
Type 4	102 100 01	2.1	2.6	11.6
	102 107 04	0.2	0.2	1.2
	101 45 01	6.8	7.6	16.6
	101 68 01	8.9	9.4	16.1
Type 5	102 102 01	17.2	19.3	16.7
	102 106 01	17.5	17.1	16.5
	102 111 01	0.2	0.5	0.6
	102 112 01	0.2	0.4	0.3

mosaic protein, 6 more specimens of genotypes 2, 3, and 4 were found anti-NS4-positive: 2 specimens from genotype 2, 3 from genotype 3, and 1 from genotype 4 (Table II). Except for the genotype 4 specimen, the NS4 mosaic protein strongly immunoreacted with these serum specimens. It is interesting to note that the weakly immunoreactive specimen was from genotype 4. As described earlier, no sequences from this genotype were included in the mosaic protein. Thus, the NS4 mosaic protein demonstrated notably more uniform immunoreactivity with serum specimens of different genotypes compared to conventional recombinant NS4 antigens (Table II).

DISCUSSION

In this study, we designed and constructed a new artificial HCV antigen composed of antigenic epitopes derived from several genotypes of the HCV NS4 protein. The strategy of making artificial antigens composed of a mosaic of small antigenic regions was employed previously for hepatitis B virus [Kumar et al., 1992], hepatitis E virus [Khudyakov et al., 1994], and HCV [Yagi et al., 1996]. The strategy has several very important and unique advantages. First, the application of mosaic antigens may improve the specificity of assay for antibody detection. It is known that 3% to 4% of antibodies that recognize an infectious agent may also immunoreact with host-specific proteins [Srinivasappa et al., 1986]. For example, molecular mimicry was found between a protein of the human immunodeficiency virus and human brain proteins [Trujillo et al., 1993]. Similarities between virus-specific antigenic epitopes and host-specific proteins were implicated in some cases of autoimmunity, the result of antibody cross-reactivity [Fujinami and Oldstone, 1989]. By carefully selecting short regions, nonspecific immunoreactivity can be reduced. Therefore, the false-positive rate would be reduced significantly. The strategy of excluding nonspecific epitopes from the design of an artificial protein was used to construct the hepatitis E mosaic antigen [Khudyakov et al., 1994]. Recently, nonspecific immunoreactivity was detected with synthetic peptides derived from region 59 [Wienhues et al., 1998]. Because the NS4 mosaic protein was constructed before this publication, sequences associated with this nonspecific immunoreactivity were included in the artificial NS4 mosaic protein. Although no false-positive results were observed in the present study, more extensive testing may reveal this problem. Research is currently in progress to identify false-positive serum specimens immunoreactive with the NS4 mosaic protein or with synthetic peptides containing different antigenic epitopes of the HCV NS4 protein. The data obtained in this study will be used to design, if necessary, new artificial proteins.

Another important advantage of the mosaic protein strategy is its flexibility in combining diagnostically relevant antigenic regions derived from different proteins from different viral strains or variants. The HCV genome is very heterogeneous [Dusheiko et al., 1994; Bukh et al., 1995; Simmonds, 1995b], and six major genotypes have been described, based on this heterogeneity [Simmonds, 1993a]. This sequence genetic variation has a significant effect on the antigenic properties of the HCV proteins [Chan et al., 1991; Machida et al., 1992; McOmish et al., 1993; 1994; Nagayama et al., 1993; Simmonds et al., 1993b; Alonso et al., 1994; Bhattacherjee et al., 1995; Zein et al., 1995; 1997; Dhaliwal et al., 1996]. Region 5-1-1 is one of the most heterogeneous antigenic regions of the HCV polyprotein [Dusheiko et al., 1994; Bukh et al., 1995; Simmonds, 1995b]. This diversity has allowed the development of serotyping assay by using synthetic peptides derived from this region [Simmonds et al., 1993b, Bhattachejee et al.,1995 ; Zhang et al., 1995]. Also, this property has been shown to contribute to a reduced immunoreactivity with serum specimens of genotypes 2–6 of various HCV diagnostic assays that are based on the genotype 1 proteins [Simmonds et al., 1993; Bhattacherjee et al., 1995; Yuki, 1995; Zein, 1995; 1997; Dhaliwal, 1996]. The artificial NS4 mosaic protein constructed in this study was designed to contain antigenic epitopes derived from 4 HCV genotypes (see the Results section). This protein detected with almost uniform efficiency, anti-HCV antibodies from serum specimens obtained from patients infected with HCV of different genotypes (Table II).

A concept similar to the hepatitis E virus mosaic protein [Khudyakov et al., 1994] and to the NS4 HCV mosaic protein was recently developed by Yagi et al. [1996]. Their artificial protein, designated as CepCM, was composed of two regions derived from the NS3 protein, and antigenic epitopes derived from core and NS4 proteins of HCV genotypes 1 and 2 [Yagi et al., 1996]. This protein demonstrated high sensitivity and specificity in detecting anti-HCV antibodies. However, no test was conducted to evaluate its efficiency in detecting anti-HCV activity in patients infected with different HCV genotypes.

Another important property of the NS4 mosaic protein is the presence of repeats of antigenic epitope variants derived from different HCV genotypes (Fig. 1). Because of this, antibodies with cross-reacting activity may potentially bind to different antigenic regions within the mosaic protein. This factor may contribute to a greater affinity of antibody binding and may explain the high sensitivity of the NS4 mosaic protein in detecting anti-HCV activity in the HCV seroconversion panels (Fig. 7), or higher endpoint titers compared to MATRIX (Fig.7).

One of the most important issues in designing and constructing artificial antigens composed of several small antigenic regions is the potential accessibility of each region for antibody binding. In the present study, we developed sequence-specific antibodies to several regions included in the NS4 mosaic protein and demonstrated that each of those regions was exposed in a manner that allowed antibody recognition. Similar results were obtained for the hepatitis E mosaic protein, where each antigenic region included in the protein

was shown to be accessible to antibodies [Khudyakov et al., 1994]. While the CepCM is an artificial protein and contains epitopes from various gene products and genotypes, the CepCM epitopes were not tested by EIA. They were, however, tested by immunoblot analysis with monoclonal antibodies [Yagi et al., 1996]. Thus, despite the fact that each of these antigens was artificially composed, there is evidence that nearly all of the antigenic regions included in these proteins were exposed to antibody binding. This finding may be explained by the fact that these proteins were designed to contain hydrophilic regions with a high density of predicted random coil and beta-turn structures. These structures are known to facilitate a loose macrostructure, which allows each epitope to be accessible to antibodies.

In conclusion, the artificial NS4 mosaic protein designed and constructed in this study was shown to be a diagnostically relevant antigen suitable for the very efficient detection of anti-NS4 activity in serum specimens. The results obtained with this protein further substantiate the concept of artificial diagnostic targets [Kumar et al., 1992, Khudyakov et al., 1994, Yagi et al., 1996].

REFERENCES

Adams NJ, Chamberlain RW, Taylor, LA, Davidson F, Lin CK, Elliott RM, Simmonds P. 1997. Complete coding sequence of hepatitis C virus genotype 6a. Biochem Biophys Res Commun 234:393–396.

Aichele P, Hengartner H, Zinkernagel RM, Schult M. 1990. Anticytoxic T-cell response induced by in vivo priming with a free synthetic peptide. J Exp Med 171:1815–1820.

Alonso C, Qu D, Lamelin JP, De Sanjose S, Vitvitski L, Li J, Berby F, Lambert V, Cortey ML, Trepo C. 1994. Serological responses to different genotypes of hepatitis C virus in France. J Clin Microbiol 32:211–212.

Barany G, Merrifield RB. 1980. The Peptides (E. Gross and J. Meienhofer, editors). New York: Academic Press, p 1–284.

Berasain C, Garcia-Granero M, Riezu-Boj J, Civeira MP, Prieto J, Borras-Cuesta F. 1993. Detection of anti-hepatitis C virus antibodies by ELISA using synthetic peptides. J Hepatol 18:80–84.

Bhattacherjee V, Prescott LE, Pike I, Rodgers B, Bell H, El-Zayadi AR, Kew MC, Conradie J, Lin CK, Marsden H, Saeed AA, Parker D, Yap P-L, Simmonds P. 1995. Use of NS-4 peptides to identify type-specific antibody to hepatitis C virus genotypes 1, 2, 3, 4, 5, and 6. J Gen Virol 76:1737–1748.

Bukh J, Purcell RH, Miller RH. 1993. At least 12 genotypes of hepatitis C virus predicted by sequence analysis of the putative E1 gene of isolates collected worldwide. Proc Natl Acad Sci USA 90: 8234–8238.

Bukh J, Miller RH, Purcell RH. 1995. Genetic heterogeneity of hepatitis C virus: quasispecies and genotypes. Semin Liver Dis 15:41–63.

Chamberlain RW, Adams N, Saeed AA, Simmonds P, Elliott RM. 1997. Complete nucleotide sequence of a type 4 hepatitis C virus variant, the predominant genotype in the middle east. J Gen Virol 78:1341–1347.

Chan SW, Simmonds P, McOmish F, Yap PL, Mitchell R, Dow B, Follett E. 1991. Serological responses to infection with three different types of hepatitis C virus. Lancet 338:1391.

Chang JC, Seidel, C, Ofenloch, B, Jue DL, Fields, HA, Khudyakov, YE. 1999. Antigenic heterogeneity of the hepatitis C virus NS4 protein as modeled with synthetic peptides. Virology 257:177–190.

Chiba J, Ohba H, Matsuura Y, Atanabe Y, Katayama T, Kikuchi S, Saito I, Miyamura T. 1991. Serodiagnosis of hepatitis C virus (HCV) infection with an HCV core protein molecularly expressed by a recombinant baculovirus. Proc Natl Acad Sci USA 88:4641–4645.

Chien DY, Choo QL, Tabrizi A, Huo C, McFarland J, Berger K, Lee C, Shuster JR, Nguyen T, Moyer DL, Tong M, Furuta S, Omata M, Tegteier G, Alter H, Schiff E, Jeffers L, Houghton M, Kuo G. 1992. Diagnosis of hepatitis C virus (HCV) infection using an immunodominat chimeric polyprotein to capture circulating antibodies: reevaluation of the role of HCV in liver disease. Proc Natl Acad Sci USA 89:10011–10015.

Choo Q, Kuo G, Weiner A, Overby L, Bradley D, Houghton M. 1989. Isolation of a cDNA clone derived from a blood-borne non-A, non-B viral hepatitis genome. Science 244:359–362.

Choo QL, Richman KH, Han JH, Berger K, Lee C, Dong C, Gallegos C, Coit D, Medina-Selby A, Barr P, Weiner AJ, Bradley DW, Kuo G, Houghton M. 1991. Genetic organization and diversity of the hepatitis C virus. Proc Natl Acad Sci USA 88:2451–2455.

Dhaliwal SK, Prescott LE, Dow BC, Davidson F, Brown H, Yap PL, Follett EAV, Simmonds P. 1996. Influence of viraemia and genotype upon serological reactivity in screening assays for antibody to hepatitis C virus. J Med Virol 48:184–190.

Dusheiko G, Schmilovitz-Weiss H, Brown D, McOmish F, Yap PL, Sherlock. 1994. Hepatitis C virus genotypes: an investigation of type-specific differences in geographic origin and disease. Hepatology 19:13–18.

Fujinami RS, Oldstone MBA. 1989. Molecular mimicry as a mechanism for virus-induced autoimmunity. Immunol Res 8:3–15.

Grakoui A, Wychowski C, Lin C, Feinstone SM, Rice C. 1993. Expression and identification of hepatitis C virus polyprotein cleavage products. J Virol 67:1385–1395.

Hijikata M, Kato N, Ootsuyama Y, Nakagawa M, Ohkoshi S, Shimotohno K. 1991. Hypervariable regions in the putative glycoprotein of hepatitis C virus. Biochem Biophys Res Commun 175:220–228.

Hosein B, Fang CT, Popovsky MA, Ye J, Zhang M, Wang CY. 1991. Improved serodiagnosis of hepatitis C virus infection with synthetic peptide antigen from capsid protein. Proc Natl Acad Sci USA 88: 3647–3651.

Houghton M, Weiner A, Han J. 1991. Molecular biology of the hepatitis C viruses: Implications for diagnosis, development and control of viral disease. Hepatology 14:381–388.

Inoue Y, Suzuki R, Matsuura Y, Harda S, Chiba J, Eatanabe Y, Saito I, Miyamura T. 1992. Expression of the amino-terminal half of the NS1 region of the hepatitis C virus genome and detection of an antibody to the expressed protein in patients with liver diseases. J Gen Virol 73:2151–2154.

Kato N, Hijikata M, Ootsuyama Y, Nakagawa M, Ohkoshi S, Sugimura T, Shimotohno K. 1990. Molecular cloning of the human hepatitis C virus genome from Japanese patients with non-A, non-B hepatitis. Proc Natl Acad Sci USA 87:9524–9528.

Khudyakov YE, Favorov MO, Khudyakov NS, Cong M, Holloway BP, Padhye N, Lambert SB, Jue DL, Field HA. 1994. Artificial mosaic protein containing antigentic epitopes of hepatitis E virus. J Virol 68:7067–7074.

Khudyakov YE, Khudyakova NS, Jue DL, Lambert SB, Fang S, Fields HA. 1995. Linear B-cell epitopes of the NS3-NS4-NS5 proteins of the hepatitis C virus as modeled with synthetic peptides. Virology 206: 666–672.

Kumar V, Bansal VJ, Rao VS, Jameel S. 1992. Hepatitist B virus envelope epitopes: gene assembly and expression in Escherichia coli of an immunological reactive novel multipleepitope polypeptide 1 (MEP-1). Gene 110:137–144.

Kuo G, Choo QL, Alter HJ, Gitnick GL, Redeker AG, Purcell RH, Miyamura T, Dienstag JL, Alter MJ, Stevens CE, Tegtmeter GE, Bonino F, Colombo M, Lee WS, Kuo C, Berger K, Shuster JR, Overby LR, Bradley DW, Houghton M. 1989. An assay for circulating antibodies to a major etiologic virus of human non-Am non-B hepatitis. Science 244:362–364.

Lok AS, Chien D, Choo QL, Chan TM, Chiu EK, Cheng IK, Houghton M, Kuo G. 1993. Antibodyresponse to core, envelope and nonstructural hepatitis C virus antigens: comparison of immunocompetent and immunosuppressed patients. Hepatology 18:497–502.

Machida A, Ohnuma H, Tsuda F, Munekata E, Tanaka T, Akahane Y, Okamoto H, Mishiro S. 1992. Two distinct subtypes of hepatitis C virus defined by antibodies directed to the putative core protein. Hepatology 16: 886–891.

McOmish F, Chan SW, Dow BC, Gillon J, Frame WD, Crawford RJ, Yap PL, Follett EAC, Simmonds P. 1993. Detection of three types of hepatitis C virus in blood donors: investigation of type-specific differences in serologic reactivity and rate of alanine aminotransferase abnormalities. Transfusion 33:7–13.

McOmish F, Yap L, Dow BC, Follett EAC, Seed C, Keller AJ, Cobain

TJ, Krusius T, Kolho E, Naukkarinen R, Lin C, Lai C, Leong S, Medgyesi GA, Hejjas M, Kiyokawa H, Fukada K, Cuypers T, Saeed AA, Al-Rasheed AM Lin M, Simmonds P. 1994. Geographical distribution of hepatitis C virus genotypes in blood donors: an international collaborative survey. J Clin Microbiol 32:884–892.

Miller RH, Purcell RH. 1990. Hepatitis C virus shares amino acid sequence similarity with pestiviruses and flaviviruses as well as members of two plant virus subgroups. Proc Natl Acad Sci USA 87:2059–2061.

Mori S, Ohkoshi S, Hijikata M, Kato N, Shimotohno K. 1992. Serodiagnostic assay of hepatitis C virus infection using viral protease expressed in *Escherichia coli.* Jpn J Cancer Res 83:264–268.

Muraiso K, Hijikata M, Ohkoshi S, Cho MJ, Kikuchi M, Kato N, Shimotohno K. 1990. A structural protein of hepatitis C virus expressed in *E. coli* facilitates accurate detection of hepatitis C virus. Biochem Biophys Res Commun 172: 511–516.

Nagayama R, Tsuda F, Okamoto H, Wang Y, Mitsui T, Tanaka T, Miyakawa Y, Mayumi M. 1993. Genotype dependence of hepatitis C virus antibodies detectable by the first-generation enzyme-linked immunobsorbent assay with C 100-3 protein. J Clin Invest 92:1529–1533.

Nasoff MS, Zebedee SL, Inchauspe G, Prince AM. 1991. Identification of an immunodominant epitope within the capsid protein of hepatitis C virus. Proc Natl Acad Sci USA 88:5462–5466.

Okamoto H, Kurai K, Okada S, Yamamoto K, Lizuka H, Tanaka T. 1992) Full-length sequence of a hepatitis C virus genome having poor homology to reported isolates: comparative study of four distinct genotypes. Virology 88:331–334.

Ptitsyn OB, Finkelstein AV. 1983. Theory of protein secondary structure and algorithm of its prediction. Biopolymers 22:15–25.

Ray R, Khanna A, Lagging LM, Meyer K, Choo QL, Ralston R, Houghton M, Becherer PR. 1994. Peptide immunogen mimicry of putative E1 glycoprotein-specific epitopes in hepatitis C virus. J Virol 68:4420–4426.

Riezu-Boj JI, Parker D, Civeira MP, Phoppard D, Corbishley TP, Camps J, Castilla A, Prieto J. 1992. Detection of hepatitis C virus antibodies with new recombinant antigens: assessment in chronic liver diseases. J Hepatol 15:309–313.

Sallberg M, Ruden U, Wahren B, Magnius LO. 1992. Immunodominant regions within the hepatitis C virus core and putative matrix proteins. J Clin Microbiol 30:1989–1994.

Sambrook J, Fritsch EF, Maniatis T. 1989. Molecular Cloning. New York: Cold Spring Harbor Laboratory Press.

Selby MJ, Choo QL, Berger K, Kuo G, Glazer E, Eckart M, Lee C, Chien D, Kuo C, Houghton M. 1993. Expression, identification and subcellular localization of the proteins. J Gen Virol 74:1103–1113.

Simmonds P. 1995a. Use of NS-4 peptides to identify type-specific antibody to hepatitis C virus genotypes 1, 2, 3, 4, 5, and 6. J Gen Virol 76:1737–1748.

Simmonds P. 1995b. Variability of hepatitis C virus. Hepatology 21: 570–583.

Simmonds P, Holmes EC, Cha T-A, Chan S-W, McOmish F, Irvine B, Beall E. 1993a. Classification of hepatitis C virus into six major genotypes and a series of subtypes by phylogenetic analysis of the NS5 region. J Gen Virol 74:2391–2399.

Simmonds P, Rose KA, Graham S, Chan S-W, McOmish F, Dow BC, Follett EAC, Yap PL, Marsden H. 1993b. Mapping of serotype-specific, immunodominant epitopes in the NS-4 region of hepatitis C virus (HCV. use of type-specific peptides to serologically differentiate infections with HCV types 1, 2, and 3. J Clin Mocrobiol 31:1493–1503.

Smith DB, Johnson KS. 1988. single-step purification of polypeptides expressed in Escherichia coli as fusions with glutathione s-transferase. Gene 67:31–40.

Srinivasappa JJ, Saegusa J, Prabhakar BS, Gentry MK, Buchmeier MJ, Wiktor TJ, Koprowski H, Oldstone MBA, Notkins AL. 1986. Molecular mimicry: frequency of reactivity of monoclonal antiviral antibodies with normal tissues. J Virol 57:397–401.

Stuyver L, Van Arnhem W, Wyseur A, Maertens G. 1994. Cloning and phylogenetic analysis of the core, E2, and NS3/NS4 regions of the hepatitis C virus type 5a. Biochem Biophys Res Commun 202: 1308–1314.

Takamizawa A, Mori C, Fuke I, Manabe S, Murakami S, Fujita J, Onishi E, Andoh T, Yoshida I, Okayama H. 1991. Structure and organization of the hepatitis C virus genome isolated from human carriers. J Virol 65:1105–1113.

Takeuchi K, Kubo Y, Boonmar Y, Waranabe Y, Katayama T, Choo QL, Kuo G, Houghton M, Saito E, Miyamura T. 1990. The putative nucleocapsid and envelope protein genes of hepatitis C virus determined by comparison of the nucleotide sequences of two isolates derived from an experimentally infected chimpanzee and healthy human carriers. J Gen Virol 71:3027–3033.

Trujillo JR, Mclane MF, Lee TH, Essex M(1993. Molecular mimicry between the human immunodeficiency virus type 1 gp120 V3 loop and human brain protein. J Virol 67:7711–7715.

Van Der Poel CL, Cuypers HT.M, Reesink HW, Weiner AJ, Quan S, Di Nello R, Van Boven JJP, Winkel I, Mulder-Folkerts,D, Exel-Oehers PJ, Schaasberg W, Leentvaar-Kuypers A, Polito A, Houghton M, Lelie PN. 1991. Confirmation of hepatitis C virus infection by new four-antigen recombinant immunoblot assay. Lancet 337: 317–319.

Weiner AL, Brauer MJ, Rosenblatt J, Richman KH, Tung J, Crawford K, Bonino F. 1991. Variable and hypervariable domains are found in the regions of HCV corresponding to the lavivirus envelope and NS1 proteins and the pestivirus envelope glycoproteins. Virology 180:842–848.

Wienhues U, Ihlenfeldt H, Seidel C, Schmitt U, Kraas W, Jung G. 1998. Characterization of a linear epitope in the nonstructural region 4 of hepatitis C virus with reactivity to seroconversion antibodies. Virology 245:281–288.

Yagi S, Kashiwakuma T, Yamaguchi K, Chiba Y, Ohtsuka E, Hasegawa A. 1996. An epitope chimeric antigen for the hepatitis C virus serological screening test. Bio Pharm Bull 19:1254–1260.

Yuki N, Hayashi N, Mita E, Hagiwara H, Oshita M, Ohkawa K, Katayama K, Kasahara A, Fusamoto H, Kamada T. 1995. Clinical characteristics and antibody profiles of chronic hepatitis C patients: relation to hepatitis C virus genotypes. J Med Virol 45:162–167.

Zein NN, Rakela J, Persin DH. 1995. Genotype-dependent serologic reactivities patients infected with hepatitis C virus in the United States. Mayo Clin Pro 70: 449–452.

Zein NN, Germer JJ, Went NK, Schimek CM, Thorvilson JN, Mitchell PS, Persing DH. 1997. Indeterminate results of the second-generation hepatitis C virus (HCV) recombinant immunoblot assay: significance of high-level c22-3 reactivity and influence of HCV genotypes. J Clin Microbiol 35:311–312.

Zhang ZX, Yun Z, Chen M, Sonnerborg A, Sallberg M. 1995. Evaluation of a multiple peptide assay for typing of antibodies to the hepatitis C virus: relation to genomic typing by the polymerase chain reaction. J Med Virol 45:50–55.